DOCUMENT RESUME

ED 329 497 SO 030 253

AUTHOR Kotowski, Jim
TITLE Critical Thinking Content in the Iowa Tests of Educational Development. Critical Thinking Reports Number 6.
INSTITUTION Illinois Univ., Urbana. Dept. of Educational Policy Studies.
PUB DATE 90
NOTE 38p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Citizenship Education; *Critical Thinking; High Schools; *Standardized Tests; Student Educational Objectives; Test Items; *Test Validity; Thinking Skills
IDENTIFIERS *Iowa Tests of Educational Development

ABSTRACT

The Iowa Test of Educational Development (ITED), a test of general abilities given to high school students, is intended as a tool for teachers and administrators to measure the progress of their students and plan accordingly; it also purports to assess the use by students of critical and evaluative abilities. Because such tests as the ITED carry significant weight in the educational system and in society at large, and because critical thinking skills have come to be recognized as one of the most important goals for education, it seemed appropriate in this report to investigate the validity of claims of the ITED to assess critical thinking ability. The overall structure of the ITED, the subtests, and procedures are analyzed and it is concluded that while the ITED is definitely better than most other tests in the area of critical thinking, it simply falls short of a comprehensive, balanced critical thinking assessment. An appendix presents, item by item, the critical thinking abilities found in the ITED. An 11-item list of references also is included. (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

JAMES F. KOTOWSKI

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SO 020 253

BEST COPY AVAILABLE

Illinois Critical Thinking Project

List of Reports

Bruce L. Stewart (1978), Teaching Logic in Grades 7-12: A Literature Review, (37 pages), Critical Thinking Reports, Number 1. ERIC document-- 183 587.

Bruce L. Stewart (1979), Testing for Critical Thinking: A Review of the Resources, (134 pages), Rational Thinking Reports, Number 2. ERIC document-- 183 588.

Thomas N. Tomko and Robert H. Ennis (1979), Evaluation of Informal Logic Competence, (53 pages), Rational Thinking Reports, Number 3. ERIC document-- 183 589.

Stephen P. Norris (1979), The Dependability of Observation Statements, (62 pages), Rational Thinking Reports, Number 4. ERIC document-- 183 590.

Michelle B. Fisher (1989), A Pilot Study on the Extent to which a Standardized Subject-Matter Test Assesses Critical Thinking Abilities, (17 pages), Critical Thinking Reports, Number 5. ERIC document-- in process, May, 1990.

Jim Kotowski (1990), Critical Thinking Content in the Iowa Tests of Educational Development, (33 pages), Critical Thinking Reports, Number 6. ERIC document-- in process, May, 1990.

Todd D. Dinkelman (1990), Critical Thinking Content in Verbal Problems of the P-SAT and P-ACT+, (57 pages), Critical Thinking Reports, Number 7. ERIC document-- in process, May, 1990.

CRITICAL THINKING CONTENT IN THE IOWA TESTS OF EDUCATIONAL DEVELOPMENT

BY

JIM KOTOWSKI

CRITICAL THINKING REPORTS

NUMBER 6

General Editor of Critical Thinking Reports
Robert H. Ennis

Published by

The Illinois Critical Thinking Project

Department of Educational Policy Studies
University of Illinois at Urbana-Champaign
Urbana, Illinois
1990

TABLE OF CONTENTS

Introduction: Why is critical thinking content in standardized tests important?.....1

Overall Structure of the ITED.....4

Subtests.....6

Procedure.....8

Results and Discussion.....10

A Possible CT Score.....17

A Caution.....18

Other Issues.....19

Conclusion.....22

Appendix.....24

References.....32

A CRITICAL THINKING SURVEY OF THE IOWA TESTS OF EDUCATIONAL DEVELOPMENT

BY JIM KOTOWSKI

INTRODUCTION: WHY IS CRITICAL THINKING CONTENT IN STANDARDIZED TESTS IMPORTANT?

For quite some time educators have argued that a major goal of schooling is that of thinking well. In 1945, the widely-quoted Harvard Committee cited three educational abilities which "should be sought above all others", one of them being "to think effectively" (p.65). In 1961, the Educational Policies Commission of the National Education Association stated that, "the purpose which runs through and strengthens all other educational purposes--the common thread of education--is the development of the ability to think." (pp.11-12). A report of The Commission on the Humanities in 1980 urged that "the Department of Education should define critical thinking as one of the basic abilities that provides the foundation for advanced abilities of all kinds." (p.37). And even more recently, in 1986, both the Holmes Group and Carnegie Task Force on Teaching as a Profession named the preparation of students and teachers to think for themselves as a primary goal.

Accompanying this continuing emphasis on thinking is an

ever-increasing reliance on standardized tests, partly as a result of the "accountability movement", wherein school-wide performance on standardized tests is published in the media. The success of teachers and of schools is measured against the ruler of test scores, for public approval or disapproval is at stake. Another effect of accountability is the pressure some families feel to locate or relocate in a high-scoring area (which presumably provides quality education). For school officials, the effect that testing thereby has on property values in their community adds to the pressure to improve scores.

With factors such as the worth of specific teachers, schools and real estate being significantly influenced by test scores, the tendency to aim education towards improving those marks can be expected to increase. If the tests on which school rank is based are going to carry such weight, then, given a need for teaching critical thinking, it is essential that the extent to which these tests assess critical thinking be investigated.

Our concern is that tests influencing huge numbers of people without paying serious attention to critical thinking will only serve to entrench an inability to think. In short, we are undertaking an examination of tests because they, if they ignore critical thinking (henceforth "CT"), may thereby promote institutionalized mediocrity.

What is at stake is not simply the claimed success or failure of our schools. The failure to think critically leaves one susceptible to deception and poor judgment on the larger, political

scale as well as on the everyday level. The Age of Mass Media has fashioned decisions at the polls, for instance, into a tension between individual deliberation and large-scale manipulation. Critical thinking, then, has become all the more essential.

In this particular report we will be looking at the Iowa Test of Educational Development (hereafter ITED), a test of general abilities given to high school students. This test is intended as a tool for teachers and administrators to measure the progress of their students and curricula and plan accordingly; "progress" is measured against a supposedly national standard. The ITED is presented as reflecting the degree to which students have developed the abilities needed to "exercise the rights and responsibilities of citizenship, the need to interpret quantitative data, and the need to find and use data from a variety of sources." (Manual for Teachers, Administrators, and Counselors, 1986, p.3). Furthermore, the tests purport to assess "student's ability to bring to bear critical and evaluative abilities" (Ibid, p.3--emphasis ours). It also claims to test for the ability to separate fact from opinion as well as for other critical thinking abilities (ibid, p.11). Moreover, our choice of the ITED was informed by its status as one of the ten most widely used standardized tests throughout Illinois high schools (1986 Illinois Report Card, Table 5, p.31). And finally, The Ninth Mental Measurement Yearbook notes that the ITED has one of "the most impressive pedigrees" of any standardized battery (Wardrop, 1985, p.727). So, taking the importance of CT

along with the status of the ITED, we found it appropriate to evaluate the latter along the lines of CT content.

It seems that by investigating the validity of claims to actual CT ability assessment on the ITED, we are breaking new ground. The Ninth Mental Measurement Yearbook, for example, takes a relatively passive stance on the content of the ITED, criticizing in some detail certain psychometric aspects while largely ignoring CT, only judging the test, as one reviewer put it, "appropriate to measure broad outcomes of schooling" (Kifer, 1985, p.723). There was no detailed analytic defense of this conclusion; the several reviewers of the ITED writing in those two editions of the Yearbook focused on the psychometric glitches , but ignored a general discussion of thinking abilities (Kifer, 1985, p.725; Wardrop, 1985, p.727).

OVERALL STRUCTURE OF THE ITED

The Iowa Tests of Educational Development (ITED) are a battery of tests designed to measure general skills, including, as mentioned above, certain critical thinking abilities. Our results show that the ITED leans in the "basic skills" direction: a couple of sections in certain subtests, almost all of certain others, and the whole of one seem to test for only low-level subject matter knowledge without recourse to CT.

The test is given to high school students, Level I for ninth and tenth graders and Level II for eleventh and twelfth. Allowance is made for exceptional students or classes, who may take the tests a level up or down from their chronological peers, depending on whether their background and/or abilities give them an advantage or disadvantage. Thus some eleventh and twelfth graders might take the Level I instead of the Level II, and other ninth and tenth-graders, if advanced enough, might take the Level II exam. The whole test takes 4 hrs. 10 min. per level and is recommended to be administered over 2-3 days.

On any given test, Levels I and II contain the same number of questions. But there is not an entirely separate group of questions for each level; they always overlap. For instance, Test L includes a total of 73 items. The first 46 make up level I, the final 46 level II; the ones in the middle belong to both. Furthermore, a "level" is not to be confused with a "part". Some subtests are divided into two sections called "parts", each with two Levels. Thus, parts 1 and 2 of Test SS are themselves each divided into Levels I and II.

The authors of the ITED indicate in the Manual, subtest by subtest, all the skills they affirm to be assessed by their exam--a loose sort of taxonomy. We can compare their taxonomy with our own, one of CT only. Through our analysis we can clarify how many of the authors' listed skills actually turn out to be CT (see "RESULTS").

SUBTESTS

The edition of the ITED evaluated here is the eighth (1987). Listed below are the subtests which comprise the test, the number of test items each contains, and the maximum time limit for completing them. The information is the same for both levels I and II. However, in the interest of evaluating as many questions as possible, we examined all items per subtest. That is, we examined the items in both levels. However, the number of items listed per test below are of one level only--the amount that a student actually taking the test would have to complete--and so are not the totals which we used to calculate percentages of CT per test.

Test E (Correctness and Appropriateness of Expression): This section has 69 items (54 for Part 1 + 15 for Part 2) to be completed in 35 minutes; it tests students' ability to recognize the best correction of commonly-made grammatical errors.

Test Q (Quantitative Thinking): There are 40 items and 50 minutes with which to complete them; this section is designed to measure students' grasp of mathematical concepts with a minimum of computation. We left Test Q out of our final analysis because it is yet unclear exactly how much

quantitative thought and CT overlap. More work is being done to resolve this issue.

Test SS (Analysis of Social Studies Materials): This subtest contains 60 items (36 in Part 1 and 42 in Part 2) to answer within 45 minutes. Part 1 consists of 24 discrete questions; Part 2 presents groups of questions, each referring to a specific passage.

Test NS (Analysis of Natural Science Materials): This test is the same length as the SS, having 60 items (24 in Part 1 and 36 in Part 2) to be completed in 45 min. Its context is scientific information, and it contains questions derived from statements, graphs, and longer passages.

Test L (Literary Materials): This is 46 items long and takes 35 minutes. It calls for interpreting, among other things, meanings of literary passages, and the gathering of evidence in favor of a certain tone, mood or meaning.

Test V (Vocabulary): There are 40 Vocabulary items to be completed within 15 minutes here. The questions provide a word in context (e.g., "<u>tepid</u> water") and five possible synonyms.

Test SI (Sources of Information): This section requires 46 items to be answered in 25 minutes. It attempts to assess students'

ability to identify the sources needed to find certain information, and to interpret the language of given reference sources.

Note: There is also a Reading Total Score derived from the 118 exercises requiring comprehension of reading selections in certain sections of different tests: Part 2 of the SS, Part 2 of the NS, and all of test L. The authors claim that these sections especially call for thinking abilities (Manual, p.23).

PROCEDURE

The classification system we used originates from Ennis' taxonomy of CT dispositions and abilities (1987, pp. 12-15), which we have revised for present purposes. It consists of two major categories, Dispositions and Abilities, each encompassing a number of aspects of critical thinking as we define it:

A. DISPOSITIONS:

1) seek a clear statement of the thesis or question
2) seek reasons
3) try to be well-informed
4) use credible sources and mention them
5) take into account the total situation
6) try to remain relevant to the main point
7) keep in mind the original and/or basic concern
8) look for alternatives
9) be open-minded
   a. Consider seriously points of view other than one's own ("dialogical thinking")

b. Withhold judgement when the evidence and reasons are insufficient

10) take a position (and change a position) when the evidence and reasons are sufficient to do so
11) seek as much precision as the subject permits
12) deal in an orderly manner with the parts of a complex whole
13) use one's critical thinking abilities
14) be sensitive to the feelings, level of knowledge, and degree of sophistication of others.

B. Abilities:

1) identifying and focusing on a question
2) analyzing the structure of arguments
3) identifying argument strategies and rhetorical techniques
4) asking and answering questions of clarification and\or challenge--filling out the details of an argument by answering questions about what is not apparent, or by having them answered
5) defining terms and judging definitions
6) employing and translating the meaning of fallacy labels
7) judging the credibility of a source
8) evaluating observations
9) deducing and judging deductions
10) induce by generalizing
11) induce by inference to best explanation
12) making and judging value judgments

Note: The induction and deduction categories do not incorporate all the instances of induction or deduction; inductive and deductive processes may be used in the context of some other ability. For example, one might induce from an argument that some assumption was made. Because of the context in which this sort of induction occurs, it is classified as "analyzing the structure of arguments" (used assumption identification, CT ability #2).

The tests were examined and scored as follows: There were three scorers, all trained in critical thinking. Scorer #1 administered subtest SS, pt.II, to himself, scoring each question according to which of the CT abilities in our taxonomy (if any) was assessed. Scorers #2 and #3 did the same. Then, all three scorers evaluated the test together, comparing their prior

findings. Where test items were scored differently by one or more of us, we discussed them until there was mutual agreement. In this way we evaluated the SS pt.II; the rest of the subtests were primarily assessed by scorer #1 in consultation with scorer #2, without a three-way discussion.

Our analysis itself is a content analysis. That is, instead of asking about what critical thinking abilities a typical student would or would not have employed in taking the ITED, we rather asked a different question: Had we designed this test item, what is the critical thinking disposition or ability we would have been attempting to assess? This seems to have been the wisest course in our situation, because if teachers today are "teaching toward the test", they will base their lessons on the abilities they see as being assessed by the test in question. The other option--ascertaining the abilities that students would actually use--would, although a valuable study, not necessarily be indicative of the material that teachers believe the test covers. Therefore we have undertaken an analysis based on content.

RESULTS AND DISCUSSION

Our analysis of the ITED was encouraging, given what we see in most other standardized tests. As noted earlier, it was more of a basic skills test than a critical thinking test. Although about half of the items assessed CT abilities, there was a

predominance of certain ones and a corresponding absence of others. For instance, induction and deduction were very common but most other abilities were touched on only sparingly, and the CT dispositions not at all.

As shown in Table I (p.12, below), the most common major CT abilities assessed were 1) induction, 2) deduction, 3) analyzing the structure of arguments, and 4) identifying argument strategies, rhetorical techniques and types of statements. The total number of questions assessing these top four was 148 out of 460, the percentage 32%. Given that the total percentage of questions assessing <u>any</u> CT ability was only three points higher at 35%, we see that the other eight abilities were paid scant, if any, attention. These findings contrast with the impression one might receive from the manual, wherein every question is assigned a certain "ability" for which it presumably tests. But their taxonomy is more general than our own, including many abilities which we do not consider to be critical thinking. In fact, almost every item on some subtests, although assigned an ability by the manual, did not call for CT.

But here we should clarify our position. Most of the non-CT abilities we found in the ITED were of the subject-matter knowledge or background knowledge variety (see below in Table II, p.13, for an explanation of these "Other Abilities"). We are not saying that they are superfluous or unnecessary; in most cases

TABLE I: # OF QUESTIONS ASSESSING SPECIFIC CT ABILITIES

	SS	NS	L	V	SI	E	TOTAL
FOCUSING ON A QUESTION	0	2	0	0	0	0	= 2
ANALYZING THE STRUCTURE OF ARGUMENTS	14	4	0	0	0	1	= 19
ID'ING ARGUMENT STRATEGIES, RHET TECHNIQUES AND TYPES OF STATEMENTS	14	1	0	0	0	0	= 15
QUESTIONS OF CLARIFICATION AND/OR CHALLENGE	0	0	0	0	0	0	= 0
DEFINING TERMS AND JUDGING DEFINITIONS	0	0	0	0	0	0	= 0
KNOW THE MEANING OF FALLACY LABELS	0	0	0	0	0	0	= 0
JUDGING CREDIBILITY	4	0	0	0	1	0	= 5
EVALUATING OBSERVATIONS	0	2	0	0	0	0	= 2
DEDUCTION	13	16	1	0	1	2	= 33
INDUCTION (GENERALIZING)	8	4	7	0	0	0	= 19
INDUCTION BY INFERENCE TO BEST EXPLANATION	17	20	23	0	1	1	= 62
MAKE AND JUDGE VALUE JUDGMENTS	2	0	0	0	0	0	= 2

	SS	NS	L	V	SI	E
QUESTIONS ASSESSING A CT ABILITY PER TOTAL # ITEMS	68/90	50/72	31/73	0/60	3/69	4/96
PERCENTAGE OF TEST QUESTIONS MEASURING SOME CT ABILITY	75%	56%	42%	0%	4.3%	4.2%

PERCENTAGE OF ITED ITEMS MEASURING A CT ABILITY= 156/460= 34%

they are simply not enough to be called CT. Furthermore, we are not implying and do not believe that CT can be done without knowledge of the topic. We hold, rather, that much tested knowledge is inert and can exist without automatically involving CT. Examples of such knowledge can be found in categories 1-5 of Table II, below.

TABLE II: Other Abilities Assessed (not critical thinking):

1) exemplification: naming\recognizing a conceptual category from examples or naming\recognizing examples given a category.
2) background knowledge: the recognition of a term or concept already known from background which does not imply CT
3) subject-matter knowledge: similar to background knowledge but arising from facts known within a specific field of study
4) translation: the recognition of a certain concept stated in different words
5) reading: the simple recognition of something's having been stated or not
6) test-wiseness: the use of process-of-elimination techniques and other clever devices to arrive at the answer

---

Following are some examples of what we did judge to be CT on the test.

Item 28 in SS assesses the CT ability "induction by inference to best explanation", asking for the hypothesis which best explains the data:

> 28. Which of the following suggests the most likely reason newspapers use cartoons to express political, social, and economic views?
>
> A To provide humor
> B To state opinions in an unbiased way
> C To provide material for non-English-speaking people
> D To portray ideas in a novel way

Answer "D" is correct because it best explains the use of cartoons in newspapers, thus, "induction by inference to best explanation".

Another item, also from the SS, is an item which, we judged, calls for deduction:

> 2. Which of the following is the least important factor for Tina to consider before deciding whether to buy a used bicycle at a garage sale?
>
> A The use she plans to make of the bicycle
> B The accuracy of the owner's statement that the bike was "like new"
> C The general quality of other merchandise at the sale
> D The price of a new bicycle

The above is a typical example of the items on the ITED that are classified as deduction. As in any case of deduction, the conclusion follows necessarily from its premises. But this one also requires an assumption to be made based on background or subject-matter knowledge: "The use one plans to make, the newness, and the price of merchandise are the most relevant concerns (while the condition of other merchandise is not as relevant) when deciding on a purchase". Within this assumption, the answer is "C" because it follows necessarily from the given information and the above assumption. Contrast #2 with an example from the Cornell Critical Thinking Test (Level X, #53, 1985):

> 53. "If these beings are from Earth, then another space ship must have landed on Nicoma. But no other space ship has landed on Nicoma."
>
> Which follows?
>
> A. Another space ship has landed on Nicoma.

B. These beings are not from Earth.
C. These beings came here by mistake.

In this instance, the conclusion follows necessarily from the premises without need of a gap-filling assumption. We present this standard example of deduction to clarify the nature of most of the deduction items on the ITED, as the need for assumptions to make the conclusions follow might obscure their deductive character.

As mentioned earlier, some other questions present on the ITED did not in our estimation assess critical thinking, but were nevertheless assigned an ability suggesting CT by the ITED Manual. One such item was from SS part I:

> 11. Speaker: "The government programs for the retired and the unemployed need strengthening; benefits need to be increased and the period of eligibility lengthened."
>
> Which of these economic goals is being emphasized by this speaker?
>
> A Economic freedom
> B Economic security
> C Full employment
> D Price stability

The answer is B, "economic security". In the Manual (p.14) the ability assigned is "Recognizing author's purpose, assumptions, and argumentative techniques", calling to mind the "identifying argument strategies, rhetorical techniques and types of statements" designation in our taxonomy. However, we judged it not to call for CT because it asks the testee only to name the "goal" already described. Naming or recognizing a conceptual category in this way is what we call "exemplification", which, although important, does

not seem to be a CT ability. It concerns us that the authors of the ITED would place a potentially misleading label on such an activity.

The critical thinking items that we did find, expressed as percentage of test items which measured to some degree at least one CT ability, follow in Table III:

TABLE III: Percentage of Questions Assessing a CT Ability

TEST SS	75%
SS part I	86%
SS part II	68%
TEST NS	56%
NS part I	61%
NS part II	52%
TEST L	42.0%
TEST SI	04.3%
TEST E	04.2%
TEST V	00.0%

The percentages in Table III are calculated using the number of questions incorporating at least one CT ability divided by the total number of questions in each test. Several items did test for more than one ability at a time; they were only counted once.

A POSSIBLE CT SCORE

The figures themselves suggest that one's total score on the ITED is not the best index of critical thinking abilities available from the ITED. However, a composite total is not the only evaluation tool used. There is also a "Reading Total" score derived from performance on 118 specific items referring to different reading selections. Reading Total items come from Part 2 of the SS, Part 2 of the NS, and all of test L . Since the authors claim that these sections especially call for thinking abilities (Manual, p.23), we investigated the possibility that this composite score could be a usable index of CT abilities.

According to our analysis, these sections, compared to the total score, do indeed assess higher percentages of CT abilities. But both SS I and NS I show even higher numbers than SS II and NS II, which are the sections from which the Reading Score is derived. Actually, both parts of the SS contain higher percentages of CT than either of the NS sections. For this reason, if there were to be an index at all, the Reading Score would not be it. Based on raw percentages, the best measures of critical thinking from the ITED would be scores on the SS I and II and NS part I, combined in the "Reading Score" fashion.

A CAUTION

The raw percentages of items which do assess CT may be misleading because they do not reflect our finding that only a few CT abilities were tested to a significant degree. Seven abilities all told, a majority of the twelve recognized by this report, combined were addressed by only 2.9 percent of ITED items. So, in addition to being a test in which no more than a third of the items require CT to any degree, the ITED leaves more than half of the CT abilities virtually untouched, and, as earlier noted, does not assess the CT dispositions at all. Keeping in mind that this is a test directed at students of high-school age, the question arises, Do we want high-school students' educational development to be assessed by an instrument that gives this amount of emphasis to critical thinking? In utilizing this test as an indicator of total educational development, school systems are not doing justice to the important goal of CT. Again, however, the ITED is a large step in the right direction, given what we have seen on most other standardized tests.

OTHER ISSUES

Up until now our commentary has focused on the percentages of items assessing CT. But other considerations require attention. First, there were four questions in the social studies section based on on an unclear distinction between fact and opinion. To qualify a statement as "opinion", it seems to use a criterion of a) unverifiability (if it cannot be proven, then it must be an opinion), or b) belief (as a belief, it is not established and therefore is an opinion). According to the fact/opinion distinctions described by Perry Weddle (1985, pp.21-22), however, even some 'facts' are unverifiable and can furthermore be framed as statements of belief just like any opinion. He criticizes these criteria and others as insufficient and confusing, and we, in turn, would have to challenge the ITED for employing them.

We did count the four questions mentioned as critical thinking in spite of this problem, noting the "fact versus opinion" issue in the appendix by the items where it surfaced.

We did not heavily employ theoretical considerations like the above in our study of the ITED. The absence of many abilities was our primary focus. But another major problem for those who wish to accurately assess critical thinking on a large scale is the absence of critical thinking dispositions from the ITED and from standardized tests in general. To possess the abilities only, while lacking such important tendencies as "Be open-minded"

(disposition 9), anyone might easily abuse her intellectual powers. At least, in having the open-minded, thorough-going attitude reflected in the dispositions, one would have the desire to look beyond her own point of view in search of available truth, the abilities coming into play thereafter. True, one might easily learn the abilities and dispositions concurrently, but we would nevertheless suggest that the dispositions embody the "spirit" of critical thinking and must not be ignored. Yet the dispositions on the whole were not assessed by the ITED.

In light of the observed trend of "teaching toward the test", we find the void left by the exclusion of the dispositions, and indeed of many CT abilities, cause for concern--especially in a relatively popular test. Indeed, it may be well-nigh impossible to assess many of the dispositions and abilities using any multiple-choice format, because the necessary interaction with other people is thereby excluded (Norris & Ennis, 1989, p.29). One suggestion to teachers who administer the ITED would be to pursue that interaction in discussing the results of the test with their students. But this solution would not give standardized results.

The universal shortcoming of multiple-choice exams, moreover, is their assessment of the ability to identify the 'right' answer rather than to formulate it. The probable lack of creativity and minimal mastery required is cause for concern; uncreative students who get by with a surface-level knowledge of critical thinking are not likely to apply their knowledge skillfully to the outside

world. They may lack the ability to "think on their feet" even in the face of apparent scholastic success.

Another problem with the ITED is that its format does not seem to encourage test takers to reach their own conclusions on important issues. On the ITED the task was almost always to identify the components of an argument without assessing its validity. For instance, on one question in Subtest L the student is asked to select an interpretation of the statement that "there is snow in the hearts of the (American) people", but is never asked to consider whether or not it is supportable. Our definition of "critical thinking" is "reasonable, reflective thinking that is focused on deciding what to believe or do." In light of this conception of CT, we feel that the ITED should go beyond the identification of what others believe or do and encourage more personal synthesis if it is truly to test for CT.

A last question is the nature of the relationship between quantitative and verbal critical thinking. Our current taxonomy is oriented towards verbal reasoning, so our analysis of the test is deficient to the extent that the test incorporates quantitative critical thinking. We are currently working on extending our taxonomy to include the quantitative element.

CONCLUSION

This type of inquiry is still young. As such our assumed conception of CT as well as our individual contributions are still relatively unchallenged. Whether they stand the test of time remains to be seen. Also, since we have yet to work out how CT operates in a quantitative context, we may have, in omitting mathematics from our analysis, misrepresented the CT assessed. Finally, we want to emphasize that our investigation of the ITED focused on content, not on the actual operations used by a population of test takers.

Of necessity some of this discussion has been negative simply because the ITED's critical thinking content is not extensive or rigorous enough, in our judgement, to fully qualify it as an educational development test. But, once again, the ITED is definitely better in the area of critical thinking than most other tests, for example, the Iowa Test of Basic Skills, the Sequential Test of Educational Progress (the subject of our last report), and many subject-matter tests. We are not dismissing the abilities it tests for as unneeded for CT; most of them are essential components of more sophisticated critical thought. The composite of the items which test for these abilities only, however, simply falls short of a comprehensive, balanced critical thinking assessment.

So three main deficits as to CT are evident in the Iowa Tests of Educational Development. One, it bypassess many CT abilities and all of the dispositions. Two, it asks subjects to distinguish between confusing conceptions of "fact" and "opinion". And three, even where CT abilities are addressed, its multiple-choice format would seem to work against a thorough assessment of them.

We recognize, however, that the ITED itself is not purely a critical thinking test, but one of educational development, a category which incorporates critical thinking. In spite of the problems we mentioned, this test does incorporate considerably more critical thinking than most other standardized tests.

APPENDIX

The appendix presents, item by item, the critical thinking abilities we found in the ITED. Sometimes a specific item raised problems (for a discussion of them see "Results and Discussion" section), in which case a phrase describing the issue is placed in parentheses next to the assigned CT ability. Questions which called for either one ability or another (depending on the approach taken) have both listed.

TEST "SS" (ANALYSIS OF SOCIAL STUDIES MATERIALS), PARTS I & II

TEST SS PART I

1. analyzing the structure of arguments (relevance)
2. deduction
3. identifying argument strategies, rhetorical techniques and types of statements (fact/opinion)
4. ----------
5. deduction
6. deduction
7. analyzing the structure of arguments (relevance)
8. induction (best explanation)
9. analyzing the structure of arguments (relevance)
10. induction (generalizing)
11. ----------
12. induction (generalizing--tables and graphs)
13. ----------
14. analyzing the structure of arguments (relevance)
15. deduction
16. deduction
17. judging the credibility of a source
18. induction (generalizing)
19. identifying argument strategies, rhetorical techniques and types of statements (fact/opinion)
20. induction (best explanation)
21. identifying argument strategies, rhetorical techniques and types of statements (verifiability)
22. making and judging value judgments
23. induction (best explanation)
24. judging the credibility of a source
25. ----------
26. judging the credibility of a source
27. ----------
28. induction (best explanation)
29. induction (best explanation)
30. induction (best explanation)
31. induction (best explanation)

32. induction (best explanation)
33. judging the credibility of a source OR induction (generalizing)
34. induction (generalizing--tables and graphs)
35. identifying argument strategies, rhetorical techniques and types of statements (verifiability)
36. identifying argument strategies, rhetorical techniques and types of statements (verifiability)

TOTAL SS PT I= 31/36= 86%

TEST SS PART II

1. deduction (fact/opinion)
2. deduction
3. identifying argument strategies, rhetorical techniques and types of statements (fact/opinion)
4. identifying argument strategies, rhetorical techniques and types of statements (verifiability)
5. ----------
6. ----------
7. ----------
8. identifying argument strategies, rhetorical techniques and types of statements
9. analyzing the structure of arguments
10. identifying argument strategies, rhetorical techniques and types of statements (verifiability)
11. induction (best explanation)
12. identifying argument strategies, rhetorical techniques and types of statements
13. analyzing arguments (identifying conclusions)
14. ----------
15. analyzing arguments (identifying conclusions)
16. ----------
17. identifying argument strategies, rhetorical techniques and types of statements
18. ----------
19. ----------
20. ----------
21. ----------
22. ----------
23. induction (best explanation)
24. induction (best explanation)
25. induction (generalizing--tables and graphs)
26. induction (best explanation)
27. induction (best explanation)
28. identifying argument strategies, rhetorical techniques and types of statements
29. ----------
30. induction (best explanation)
31. ----------

32.induction (best explanation)
33.analyzing the structure of arguments (used assumption identification)
34.analyzing the structure of arguments (used assumption identification)
35.making and judging value judgments, induction (generalizing)
36.----------
37.deduction
38.deduction
39.deduction
40.deduction
41.----------
42.deduction
43.----------
44.deduction
45.analyzing the structure of arguments (used assumption identification)
46.----------
47.induction (best explanation)
48.----------
49.analyzing the structure of arguments (identifying conclusions)
50.induction (generalizing and best explanation)
51.induction (best explanation)
52.analyzing the structure of arguments
53.identifying argument strategies, rhetorical techniques and types of statements
54.analyzing the structure of arguments (similarities and differences)

TOTAL SS PT II= 37/54= 68%

TOTAL SS I & II= 68/90= 75%

TEST "NS" (NATURAL SCIENCES), PARTS I & II

TEST NS PART I

1.----------
2.----------
3.deduction, induction (best explanation)
4.deduction
5.----------
6.induction (planning experiments)
7.----------
8.----------
9.----------
10.deduction
11.deduction
12.----------
13.----------

14.induction (best explanation)
15.deduction
16.----------
17.deduction
18.----------
19.deduction
20.deduction-in-induction (generalizing)
21.----------
22.induction (best explanation--planning experiments)
23.induction (best explanation)
24.identifying argument strategies, rhetorical techniques, and types of statements (verifiability)
25.induction (best explanation)
26.induction (best explanation--planning experiments)
27.deduction
28.deduction
29.----------
30.----------
31.----------
32.deduction
33.induction (best explanation)
34.analyzing arguments (used assumption identification)
35.induction (best explanation)
36.evaluating observations

TOTAL NS PT I= 22/36= 61%

TEST NS PART II

1.evaluating observations
2.----------
3.----------
4.induction (best explanation--planning experiments)
5.----------
6.----------
7.focusing on a question
8.----------
9.----------
10.identification of argument strategies, rhetorical techniques and types of statements
11.----------
12.----------
13.----------
14.deduction
15.induction (best explanation)
16.----------
17.----------
18.deduction OR induction (best explanation)
19.focusing on a question
20.induction (best explanation)
21.induction (best explanation)

22.induction (best explanation)
23.----------
24.----------
25.----------
26.induction (best explanation--planning experiments)
27.induction (best explanation)
28.deduction
29.----------
30.----------
31.----------
32.induction (generalizing from tables and graphs)
33.induction (best explanation)
34.----------
35.----------
36.induction (best explanation)
37.induction (best explanation)
38.----------
39.----------
40.deduction
41.analyzing the structure of arguments (used assumption identification)
42.----------
43.analyzing the structure of arguments (identifying conclusions)
44.deduction
45.deduction
46.induction (best explanation)
47.induction (best explanation)
48.induction (generalizing)
49.induction (generalizing--tables and graphs)
50.----------
51.analyzing the structure of arguments (used assumption identification)
52.----------
53.----------
54.----------

TOTAL NS PT II= 28/54= 52%

TOTAL NS I & II= 50/90= 56%

TEST "L" (ABILITY TO INTERPRET LITERARY MATERIALS)

1.induction (best explanation)
2.----------
3.induction (best explanation)
4.deduction
5.induction (best explanation)
6.----------
7.----------
8.----------
9.induction (best explanation)

10.---------
11.---------
12.---------
13.---------
14.induction (best explanation)
15.---------
16.---------
17.---------
18.induction (best explanation)
19.---------
20.---------
21.---------
22.induction (best explanation)
23.---------
24.---------
25.---------
26.induction (best explanation)
27.induction (generalizing)
28.---------
29.induction (best explanation)
30.induction (best explanation)
31.induction (best explanation)
32.induction (best explanation)
33.induction (generalizing)
34.---------
35.induction (best explanation)
36.---------
37.induction (best explanation)
38.---------
39.induction (generalizing)
40.induction (best explanation)
41.---------
42.---------
43.induction (best explanation)
44.---------
45.---------
46.induction (best explanation)
47.---------
48.---------
49.---------
50.induction (best explanation)
51.---------
52.---------
53.---------
54.---------
55.induction (best explanation)
56.---------
57.---------
58.induction (best explanation)
59.analyzing the structure of arguments (identifying conclusions)
60.induction (best explanation)
61.induction (generalizing)

62.induction (generalizing)
63.---------
64.induction (generalizing)
65.---------
66.---------
67.induction (generalizing)
68.---------
69.---------
70.---------
71.induction (best explanation)
72.---------
73.---------

TOTAL L= 31/73= 42.0%

APPENDIX: TEST "V" (VOCABULARY)

All the questions in Test V called for either subject matter knowledge ("I know what this word means") or test wiseness ("I can guess"). No critical thinking was involved in this process.

TOTAL V= 0/60= 0%

TEST "SI" (USE OF SOURCES OF INFORMATION)

1-12.----------
13.induction (best explanation)
14-34.---------
35.judging the credibility of a source
36-52.---------
53.deduction
54-69.---------

TOTAL SI= 3/69= 04.3%

TEST "E" (CORRECTNESS AND APPROPRIATENESS OF EXPRESSION)

1-43.----------
44.induction (best explanation)
45.----------
46.deduction
47-52.---------
53.deduction
54.analyzing the structure of arguments (identifying conclusions)
55-96.---------

TOTAL E= 4/96= 04.2%

TEST "Q" (ABILITY TO DO QUANTITATIVE THINKING)

We decided to eliminate Test Q from our analysis of critical thinking content because, in general, its math problems seem to require a different type of thinking than those to which we can apply our taxonomy.

References

Buros Institute of Mental Measurements (1985). The Ninth Mental Measurements Yearbook. Lincoln, NA: University of Nebraska Press.

Commission on the Humanities (1980). The humanities in American life. Berkeley, CA: University of California Press.

Educational Policies Commission (1961). The central purpose of American education. Washington, D.C.: National Educational Association.

Ennis, R. H. & Millman, Jason (1985). Cornell Critical Thinking Test Level X (p.10).Pacific Grove, CA: Midwest Publications.

Ennis, R. H. (1987b). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & R.J. Sternberg (Eds.), Teaching thinking skills (pp.9-26). New York: Freeman.

Harvard Committee (1945). General education in a free society. Cambridge, MA: Harvard University Press.

Holmes Group (1986). Tomorrow's teachers. East Lansing, MI: author.

Illinois State Board of Education (1987). Performance profiles: Illinois Schools Report to the public; a summary and analysis of information from the 1986 report cards. Springfield, IL: Illinois State Board of Education.

Iowa Tests of Educational Development (1986). Manual for Teachers, Administrators, and Counselors (pp.3-23). Iowa City, Iowa: The University of Iowa.

Norris, Stephen P. & Ennis, R.H. (1989). Evaluating Critical Thinking (p.29). Pacific Grove, CA: Midwest Publications.

Weddle, Perry (1985). Fact from opinion. Informal Logic, VII.1, 19-26.